OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: November 17, 2003, 20:18:19 ; Search time 2564 Seconds
(without alignments)
16626.292 Million cell updates/sec

Title: US-09-894-633A-81
Perfect score: 666
Sequence: 1 ctgcggctgctatccaccact........aacggccggcggggcgtggg 666

Scoring table: OLIGO_NJC
Gapop 60.0 , Gapext 60.0
Searched: 2888711 seqs, 2045813386 residues
Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pri:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vi:*
29: em_vr:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:**

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Query Length	DB	ID	Description
	1	666	100.0	666	6	AX356287	AX356287 Sequence
	2	112	16.8	7142	8	AF451895	AF451895 Zea mays
c	3	30	4.5	127652	2	AP001526	AP001526 Homo sapi
c	4	30	4.5	132948	2	AC006410	AC006410 Homo sapi
	5	30	4.5	154841	2	AC090384	AC090384 Homo sapi
c	6	30	4.5	155736	2	AC004229	AC004229 Homo sapi
	7	26	3.9	1092	6	AX653891	AX653891 Sequence
	8	26	3.9	90837	10	AL928648	AL928648 Mouse DNA
	9	26	3.9	120926	2	AC133008	AC133008 Oryza sat
	10	26	3.9	155185	2	AC023549	AC023549 Homo sapi
	11	26	3.9	157834	2	AC022931	AC022931 Homo sapi
	12	26	3.9	164526	9	AC114969	AC114969 Homo sapi
c	13	26	3.9	183011	2	BX293990	BX293990 Mus muscu
c	14	26	3.9	190338	10	AC129216	AC129216 Mus muscu
c	15	26	3.9	199873	10	AL928914	AL928914 Mouse DNA
c	16	26	3.9	206660	10	AC122356	AC122356 Mus muscu
	17	26	3.9	212597	2	AC105463	AC105463 Rattus no
c	18	26	3.9	221389	2	AC112122	AC112122 Rattus no
c	19	26	3.9	221951	10	AC122857	AC122857 Mus muscu
c	20	26	3.9	235511	2	AC121887	AC121887 Mus muscu
	21	26	3.9	255090	2	AC098334	AC098334 Rattus no
c	22	26	3.9	265405	2	AC099437	AC099437 Rattus no
	23	26	3.9	273204	2	AC119488	AC119488 Rattus no
c	24	26	3.9	275319	2	AC109121	AC109121 Rattus no
	25	26	3.9	320828	2	AC098399	AC098399 Rattus no
	26	25	3.8	480	14	SHU27486	U27486 Suid herpes
c	27	25	3.8	110753	9	AC010216	AC010216 Homo sapi
c	28	25	3.8	118436	8	AP005064	AP005064 Oryza sat
c	29	25	3.8	142986	2	AC023467	AC023467 Homo sapi
	30	25	3.8	163584	2	AC008470	AC008470 Homo sapi
	31	25	3.8	198911	2	AC111975	AC111975 Rattus no
	32	25	3.8	225379	2	AC121485	AC121485 Rattus no
	33	25	3.8	248933	2	AC109773	AC109773 Rattus no
c	34	25	3.8	249105	2	AC094445	AC094445 Rattus no
c	35	25	3.8	250169	2	AC120668	AC120668 Rattus no
	36	25	3.8	261502	2	AC094923	AC094923 Rattus no
	37	24	3.6	552	2	AY113570	AY113570 Drosophil
	38	24	3.6	1923	8	BNU59446	U59446 Brassica na
c	39	24	3.6	3200	3	AY095032	AY095032 Drosophi
c	40	24	3.6	6062	2	AF139177	AF139177 Streptomy
	41	24	3.6	13002	2	AL137126	AL137126 Continuation 4/5 of
	42	24	3.6	17528	2	AC020192	AC020192 Drosophil
	43	24	3.6	31741	9	AL607065	AL607065 Human DNA
c	44	24	3.6	46630	2	DMC115C2	DMC115C2 Drosophil
c	45	24	3.6	60135	2	AC091200	AC091200 Drosophil

ALIGNMENTS

RESULT 1
AX356287
LOCUS   AX356287   666 bp   DNA   linear   PAT 06-FEB-2002
DEFINITION Sequence 81 from Patent WO0200905.
ACCESSION   AX356287
VERSION   AX356287.1 GI:18620794
KEYWORDS
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Conner,T.W., Dubois,P., Malven,M. and Masucci,J.D.
  TITLE   Plant regulatory sequences for selective control of gene expression

Pred. No. is the number of results predicted by chance to have a

```
JOURNAL     Patent: WO 0200905-A 81 03-JAN-2002;
            Monsanto Technology LLC (US)
FEATURES    Location/Qualifiers
  source    1..666
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /db_xref="taxon:4577"
BASE COUNT  95 a       272 c      159 g      140 t
ORIGIN

Query Match     100.0%;   Score 666;  DB 6;  Length 666;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  666;  Conservative   0;  Mismatches    0;  Indels   0;  Gaps   0;

Qy     1: CTGCGGCTGCTATCACCACTCCACCAACTCCAATTAAAGATCTCGCTACGCCAGCTTTGCA   60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1: CTGCGGCTGCTATCACCACTCCACCAACTCCAATTAAAGATCTCGCTACGCCAGCTTTGCA   60

Qy    61: CTTCCTCCTCTCTCTCCTACTGGCCGCCGGCCCTGCAGTGGCAGCGGCCCGGTTTGCTGCTAG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61: CTTCCTCCTCTCTCTCCTACTGGCCGCCGGCCCTGCAGTGGCAGCGGCCCGGTTTGCTGCTAG  120

Qy   121: CTGAGCTGCGGGGGTCGCTCTTAGATATAGCCCAAAACTCACTCCACCACTCGTTCC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121: CTGAGCTGCGGGGGTCGCTCTTAGATATAGCCCAAAACTCACTCCACCACTCGTTCC  180

Qy   181: ATGGAACCTAGACCAAAAGTACTTCACGCTCTCGCCTCTCGCCTCTCGTTCCCCTCT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181: ATGGAACCTAGACCAAAAGTACTTCACGCTCTCGCCTCTCGCCTCTCGTTCCCCTCT  240

Qy   241: CCCTCTCCGCAGCAGCAAAAGAGATCCCGGCCGAGAAGGGCGGCGGCTAGCTGCCCGGCTAGC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241: CCCTCTCCGCAGCAGCAAAAGAGATCCCGGCCGAGAAGGGCGGCGGCTAGCTGCCCGGCTAGC  300

Qy   301: TGCTGGCGCCCGGCGCCCCGGCGCCATATATCTGTCATCGCCATCACCCACCATGGCC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301: TGCTGGCGCCCGGCGCCCCGGCGCCATATATCTGTCATCGCCATCATCACCCACCATGGCC  360

Qy   361: CGGCCGGCCAACACCGCGTATTAGCTCTGTGTGTGTCGCTCGCTCCGTCCACCTGCGACCGACTGA  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361: CGGCCGGCCAACACCGCGTATTAGCTCTGTGTGTGTCGCTCGCTCCGTCCACCTGCGACCGACTGA  420

Qy   421: GGATCCATCTCCACCGAGCTCTCGCGCTCGGCTCTCCGCGCGTCCTCCTTGCGCGTCCTCCCCTC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421: GGATCCATCTCCACCGAGCTCTCGCGCTCGGCTCTCCGCGCGTCCTCCTTGCGCGTCCTCCCCTC  480

Qy   481: CGTCCCTCACGGCATCCATTTCCGTGTGCTCGGTGTGTCGCGCGGCGGTCCTCGCT  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481: CGTCCCTCACGGCATCCATTTCCGTGTGCTCGGTGTGTCGCGCGGCGGTCCTCGCT  540

Qy   541: CCGCTCTCGCTCCCTCGGCCCCCTCCTCCCCCGGCGGCCCGGACCTGCACCTGCAC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541: CCGCTCTCGCTCCCTCGGCCCCCTCCTCCCCCGGCGGCCCGGACCTGCACCTGCAC  600

Qy   601: AGATCGGGCGCGGGCGGGCCGGCCGACCGATCGATTATCGTGTCAACGGCGCGGGGGGG  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601: AGATCGGGCGCGGGCGGGCCGGCCGACCGATCGATTATCGTGTCAACGGCGCGGGGGGG  660

Qy   661: CGTGGG  666
            ||||||
Db   661: CGTGGG  666

RESULT 2
AF451895
LOCUS       AF451895       7142 bp    DNA     linear   PLN 01-MAY-2002
DEFINITION  Zea mays LIGULELESS1 (lgl1) gene, complete cds.
ACCESSION   AF451895
VERSION     AF451895.1  GI:17940733
KEYWORDS    Zea mays
SOURCE      Zea mays
  ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 7142)
  AUTHORS   Kaplinsky,N.J., Braun,D.M., Penterman,J., Goff,S.A. and Freeling,M.
  TITLE     Utility and distribution of conserved noncoding sequences in the
            grasses
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (9), 6147-6151 (2002)
  MEDLINE   21980591
  PUBMED    11972021
REFERENCE   2  (bases 1 to 7142)
  AUTHORS   Braun,D.M., Kaplinsky,N.J., Penterman,J., Goff,S.A. and Freeling,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-NOV-2001) Plant and Microbial Biology, UC-Berkeley,
            111 Koshland Hall, Berkeley, CA 94720, USA
FEATURES    Location/Qualifiers
  source    1..7142
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /db_xref="taxon:4577"
            /chromosome="2"
  gene      <3441..>6758
            /gene="lgl"
  mRNA      join(<3441..4135,6029..6183,6409..>6758)
            /gene="lgl"
            /product="LIGULELESS1"
  CDS       join(3441..4135,6029..6183,6409..6758)
            /gene="lgl"
            /note="transcription factor"
            /codon_start=1
            /product="LIGULELESS1"
            /protein_id="AAL49746.1"
            /db_xref="GI:17940734"
            /translation="MMNLSAAANGRDEFPYVVPSNAAAPPPSLLPTYEQQCSSIHR
            EHHQLGYNLEANSLALLPPSNAAAAHHTTPAGGHSPHDILHFYTPPSAASHYLA
            AAGNPYSHLVSAPGTFHQTSSSYPPAAAAQAAPEYYFPTLVSSAEENMASFAATQL
            GLNLGYRTYPPRGGYTYGHHPPRCQAESCKADLLSAKRYHRRHKVCEHHSKAPVVVT
            AGGLHQRFCQQCSRFHLLDEFDDAKKSCKRLADHNRRRRKSKFSDADAGDKKRAHAN
            KAAAAKDKAESSSKMDIGDLGAQILGSALLSKEQDQTMDLGEVVKEAVDPKGKASM
            QQHYGFPFHSSSAGSCFPQTQAVSSDTTSNIGQYQEPSLGFHHQHHQHSNILQLGQAM
            FDLDFDH"

BASE COUNT   1865 a   1882 c   1494 g   1901 t
ORIGIN

Query Match         16.8%;   Score 112;  DB 8;  Length 7142;
Best Local Similarity   100.0%;  Pred. No. 3.5e-48;
Matches   112;  Conservative   0;  Mismatches    0;  Indels   0;  Gaps   0;

Qy   315: CGCCCGCGCATATATTGTGTCATGGCCATCACCACACCATGGCCGGCCCGGCCCAACAC  374
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3077: CGCCCGGCGCATATATCTGTGTCATCGCCATCATCACCCACCATGGCCCGGCCCGGCCCAACAC 3136

Qy   375: CGCCGTATTAGCTCTGTCTGTCGCTCGCTCCGTCCACCTGCGACCTGAGCGATC  426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3137: CGCCGGTATTAGCTCTGTCTGTCGCTCGCTCCGTCCACCTGCGACCTGAGCGATC 3188

RESULT 3
AP001526/c
LOCUS       AP001526      127652 bp    DNA     linear   HTG 30-MAY-2000
DEFINITION  Homo sapiens chromosome 11 clone RP11-727C13 map 11q13, WORKING
            DRAFT SEQUENCE, 24 unordered pieces.
ACCESSION   AP001526
VERSION     AP001526.2  GI:8117367
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 127652)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens 127,652 genomic DNA of 11q13
  JOURNAL   Published Only in Database (2000)
```

```
REFERENCE   2  (bases 1 to 127652)
AUTHORS     Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       Direct Submission
JOURNAL     Submitted (22-MAR-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ. 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp,
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924)
COMMENT     On May 30, 2000 this sequence version replaced gi:7328540.
            -----------  Genome Center
               Center: RIKEN Genomic Sciences Center(GSC)
               Center code: RIKEN
               Web site: http://hgp.gsc.riken.go.jp/
               Contact: hattori@gsc.riken.go.jp
            -----------  Project Information
               Center project name: HumDraft11
               Center clone name: RP11-727C13
            -----------  Summary Statistics
               Sequencing vector: PCR products; 100% of reads
               Chemistry: Dye-terminator ET-amersham; 100% of reads
               Assembly program: Phrap; version 0.990329
               Consensus quality: 115572 bases at least Q40
               Consensus quality: 121432 bases at least Q30
               Consensus quality: 123817 bases at least Q20
               Insert size: 125352; sum-of-contigs
               Quality coverage: 6.02x in Q20 bases; sum-of-contigs
            -----------
            NOTE: This is a 'working draft' sequence. It currently consists of
            24 contigs. The true order of the pieces is not known and their
            order in this sequence record is arbitrary. Gaps between the
            contigs are represented as runs N, but the exact sizes of the gaps
            are unknown. This record will be updated with the finished sequence
            as soon as it is available and the accession number will be
            preserved
                  1          12210  contig of  12210 bp in length
              12311          30914  contig of  18604 bp in length
              31015          46674  contig of  15660 bp in length
              46775          56057  contig of   9283 bp in length
              56158          66930  contig of  10773 bp in length
              67031          73289  contig of   6259 bp in length
              73390          80253  contig of   6864 bp in length
              80354          85307  contig of   4954 bp in length
              85408          88728  contig of   3321 bp in length
              88829          90367  contig of   1539 bp in length
              90468          92766  contig of   2299 bp in length
              92867          96711  contig of   3845 bp in length
              96812         101444  contig of   4633 bp in length
             101545         105387  contig of   3843 bp in length
             105488         108838  contig of   3351 bp in length
             108939         112138  contig of   3200 bp in length
             112239         114237  contig of   1999 bp in length
             114338         117279  contig of   2942 bp in length
             117380         119677  contig of   2298 bp in length
             119778         122392  contig of   2615 bp in length
             122493         123862  contig of   1370 bp in length
             123963         125178  contig of   1216 bp in length
             125279         126524  contig of   1246 bp in length
             126625         127652  contig of   1028 bp in length
            Sequence updated (26-May-2000)
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 24 contigs. The true order of the pieces of the record
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the contigs are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    12210: contig of 12210 bp in length
            *    12211    12310: gap of 100 bp
            *    12311    30914: contig of 18604 bp in length
            *    30915    31014: gap of 100 bp
```

```
*    31015     46674: contig of 15660 bp in length
*    46675     46774: gap of 100 bp
*    46775     56057: contig of 9283 bp in length
*    56058     56157: gap of 100 bp
*    56158     66930: contig of 10773 bp in length
*    66931     67030: gap of 100 bp
*    67031     73289: contig of 6259 bp in length
*    73290     73389: gap of 100 bp
*    73390     80253: contig of 6864 bp in length
*    80254     80353: gap of 100 bp
*    80354     85307: contig of 4954 bp in length
*    85408     85407: gap of 100 bp
*    85408     88728: contig of 3321 bp in length
*    88729     88828: gap of 100 bp
*    88829     90367: contig of 1539 bp in length
*    90368     90467: gap of 100 bp
*    90468     92766: contig of 2299 bp in length
*    92767     92866: gap of 100 bp
*    92867     96711: contig of 3845 bp in length
*    96812     96811: gap of 100 bp
*    96812    101444: contig of 4633 bp in length
*   101445    101544: gap of 100 bp
*   101545    105387: contig of 3843 bp in length
*   105388    105487: gap of 100 bp
*   105488    108838: contig of 3351 bp in length
*   108839    108938: gap of 100 bp
*   108939    112138: contig of 3200 bp in length
*   112139    112238: gap of 100 bp
*   112239    114237: contig of 1999 bp in length
*   114238    114337: gap of 100 bp
*   114338    117279: contig of 2942 bp in length
*   117280    117379: gap of 100 bp
*   117380    119677: contig of 2298 bp in length
*   119678    119777: gap of 100 bp
*   119778    122392: contig of 2615 bp in length
*   122393    122492: gap of 100 bp
*   122493    123862: contig of 1370 bp in length
*   123863    123962: gap of 100 bp
*   123963    125178: contig of 1216 bp in length
*   125179    125278: gap of 100 bp
*   125279    126524: contig of 1246 bp in length
*   126525    126624: gap of 100 bp
*   126625    127652: contig of 1028 bp in length
FEATURES             Location/Qualifiers
     source          1..127652
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q13"
                     /clone="RP11-727C13"
     misc_feature    1..12210
                     /note="assembly_fragment"
     misc_feature    12311..30914
                     /note="assembly_fragment"
     misc_feature    31015..46674
                     /note="assembly_fragment"
     misc_feature    80354..85307
                     /note="assembly_fragment clone_end:T7 vector_side:left"
     misc_feature    46775..56057
                     /note="assembly_fragment"
     misc_feature    85408..88728
                     /note="assembly_fragment"
     misc_feature    56158..66930
                     /note="assembly_fragment"
     misc_feature    67031..73289
                     /note="assembly_fragment"
     misc_feature    73390..80253
                     /note="assembly_fragment"
     misc_feature    80354..85307
                     /note="assembly_fragment"
     misc_feature    88829..90367
                     /note="assembly_fragment clone_end:SP6 vector_side:right"
     misc_feature    90468..92766
                     /note="assembly_fragment"
```

```
misc_feature    92867..96711
                /note="assembly_fragment"
misc_feature    96812..101444
                /note="assembly_fragment"
misc_feature    101545..105387
                /note="assembly_fragment"
misc_feature    105488..108838
                /note="assembly_fragment"
misc_feature    108939..112218
                /note="assembly_fragment"
misc_feature    112239..114237
                /note="assembly_fragment"
misc_feature    114338..117279
                /note="assembly_fragment"
misc_feature    117380..119677
                /note="assembly_fragment"
misc_feature    119778..122392
                /note="assembly_fragment"
misc_feature    122493..123862
                /note="assembly_fragment"
misc_feature    123963..125178
                /note="assembly_fragment"
misc_feature    125279..126524
                /note="assembly_fragment"
misc_feature    126625..127652
                /note="assembly_fragment"

BASE COUNT    32152 a    31281 c    30169 g    31749 t    2301 others
ORIGIN

Query Match
Best Local Similarity    4.5%;    Score 30;    DB 2;    Pred. No. 0.00019;    Length 127652;
Matches    30;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy    219 CTCGCTCTCGCCCTCTCCCTCTCCCTCTCC 248
             ||||||||||||||||||||||||||||||
Db    79313 CTCGCTCTCGCCCTCTCCCTCTCCCTCTCC 79284

RESULT 4
AC006410/C
LOCUS    AC006410    132948 bp    DNA    linear    HTG 11-JUN-2002
DEFINITION    Homo sapiens chromosome 11 clone PAC pDJ591L21 map 11q12.2, ***
             SEQUENCING IN PROGRESS ***, 6 unordered pieces.
ACCESSION    AC006410
VERSION    AC006410.2    GI:21389726
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE    Homo sapiens (human)
ORGANISM    Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 132948)
AUTHORS    Evans,G.A., Athanasiou,M., Aguayo,P., Arenare,S., Armstrong,D.,
          Basil,M., Buettner,J., Butler,C., Card,P., deSailboat,C., Dunn,J.,
          English,C., Ethridge,S., Garner,H.R., Gordon,M., Grant,O.,
          Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T.,
          McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., and
          Schultz,R.A., Stimson,S., Wagner,N., Waller,M., Ward,T. and
          Whitaker,J.
TITLE    HTGS Submission
JOURNAL    Unpublished
REFERENCE    2 (bases 1 to 132948)
AUTHORS    Evans,G.A., Athanasiou,M., Aguayo,P., Arenare,S., Armstrong,D.,
          Basil,M., Buettner,J., Butler,C., Card,P., deSailboat,C., Dunn,J.,
          English,C., Ethridge,S., Garner,H.R., Gordon,M., Grant,C.,
          Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T.,
          McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., and
          Schultz,R.A., Stimson,S., Wagner,N., Waller,K., Ward,T. and
          Whitaker,J.
TITLE    Direct Submission
JOURNAL    Submitted (16-JAN-1999) Genome Science & Technology Center,
         University of Texas Southwestern Medical Center, 5323 Harry Hines
         Blvd, Dallas, TX 75235-8591, USA
COMMENT    On Jun 11, 2002 this sequence version replaced gi:4160136.


CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11q12.2
Best's disease region mapped between STS D11S461 and EST AHNAK.
This region spans over 1.5 Mbp.
MARKER CONFIRMATION: WI-17055
MAPPED CLONE OVERLAP: PACs pDJ57L14 and pDJ632C8.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

           1     2186: contig of 2186 bp in length
        2187     2286: gap of unknown length
        2287    12644: contig of 10358 bp in length
       12645    12744: gap of unknown length
       12745    18996: contig of 6242 bp in length
       18997    19086: gap of unknown length
       19087    35290: contig of 16204 bp in length
       35291    35390: gap of unknown length
       35391    66291: contig of 30901 bp in length
       66292    66391: gap of unknown length
       66392   132948: contig of 66557 bp in length

FEATURES    Location/Qualifiers
source       1..132948
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /chromosome="11"
             /map="11q12.2"
             /db_xref="taxon:9606"
             /clone="PAC pDJ591L21"

BASE COUNT    34604 a    34236 c    31860 g    31745 t    503 others
ORIGIN

Query Match
Best Local Similarity    4.5%;    Score 30;    DB 2;    Pred. No. 0.00019;    Length 132948;
Matches    30;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy    219 CTCGCTCTCGCCCTCTCCCTCTCCCTCTCC 248
             ||||||||||||||||||||||||||||||
Db    116441 CTCGCTCTCGCCCTCTCCCTCTCCCTCTCC 116412

RESULT 5
AC090384
LOCUS    AC090384    154841 bp    DNA    linear    HTG 26-AUG-2001
DEFINITION    Homo sapiens chromosome 11 clone RP11-727C3 map 11, WORKING DRAFT
             SEQUENCE, 13 unordered pieces.
ACCESSION    AC090384
VERSION    AC090384.3    GI:15290875
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE    Homo sapiens (human)
ORGANISM    Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 154841)
AUTHORS    Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE    Homo sapiens chromosome 11, clone RP11-727C13
JOURNAL    Unpublished
REFERENCE    2 (bases 1 to 154841)
AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
          Barna,N., Bastien,V., Boguslavkiy,L., Bockhgalter,B., Brown,A.,
          Cacarata,J., Campopiano,A., Choepel,K., Colangelo,M., Collins,S.,
          Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,C.S.,
          Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
          Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
          Hagos,B., Hexford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
          Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
          Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
          Marquis,E., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
          McPheeters,R., Meldrim,J., Mereus,L., Minova,L., Mionga,V.,
          Murphy,T., Naylor,J., Nguyen,C., Nortu,C., Norman,C.H.,
```

```
              O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
              Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
              Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
              Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
              Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
              Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
              Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
              Wilson,B., Wu.X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
              Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
  COMMENT   On Aug 26, 2001 this sequence version replaced gi:14595869.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            --------------- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence submissions@genome.wi.mit.edu
            --------------- Project Information
              Center project name: L12682
              Center clone name: 727_C13
            --------------- Summary Statistics
              Sequencing vector: Plasmid; n/a; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 150768 bases at least Q40
              Consensus quality: 152390 bases at least Q30
              Consensus quality: 153137 bases at least Q20
              Insert size: 176000; agarose-fp
              Insert size: 153641; sum-of-contigs
              Quality coverage: 10.9 in Q20 bases; agarose-fp
              Quality coverage: 12.5 in Q20 bases; sum-of-contigs
 *          NOTE: This is a 'working draft' sequence. It currently
 *          consists of 23 contigs. The true order of the pieces
 *          is not known and their order in this sequence record is
 *          arbitrary. Gaps between the contigs are represented as
 *          runs of N, but the exact sizes of the gaps are unknown.
 *          This record will be updated with the finished sequence
 *          as soon as it is available and the accession number will
 *          be preserved.
 *              1: contig of 1365 bp in length
 *           1366: gap of 100 bp
 *           1466: contig of 929 bp in length
 *           2394: gap of 100 bp
 *           2495: contig of 1336 bp in length
 *           3830: gap of 100 bp
 *           3931: contig of 4145 bp in length
 *           8075: gap of 100 bp
 *           8176: contig of 3563 bp in length
 *          11738: gap of 100 bp
 *          11839: contig of 4732 bp in length
 *          16570: gap of 100 bp
 *          16571: contig of 4281 bp in length
 *          20951: gap of 100 bp
 *          20952: contig of 6591 bp in length
 *          27642: gap of 100 bp
 *          27743: contig of 19561 bp in length
 *          47303: gap of 100 bp
 *          47404: contig of 25771 bp in length
 *          73174: gap of 100 bp
 *          73275: contig of 15795 bp in length
 *          89069: gap of 100 bp
 *          89170: contig of 33626 bp in length
 *         122795: gap of 100 bp
 *         122896: contig of 31946 bp in length
 *         154841: contig of 154841
  FEATURES  Location/Qualifiers
  source    1..154841
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
```

```
            misc_feature     /db_xref="taxon:9606"
                             /chromosome="11"
                             /map="11"
                             /clone="RP11-727C13"
                             /clone_lib="RPCI-11 Human Male BAC"
                             1..1365
                             /note="assembly_fragment"
                             clone_end:SP6
                             vector_side:left"
            misc_feature     1466..2394
                             /note="assembly_fragment"
            misc_feature     2495..3830
                             /note="assembly_fragment"
            misc_feature     3931..8075
                             /note="assembly_fragment"
            misc_feature     8176..11738
                             /note="assembly_fragment"
            misc_feature     11839..16570
                             /note="assembly_fragment"
            misc_feature     16671..20951
                             /note="assembly_fragment"
            misc_feature     21052..27642
                             /note="assembly_fragment"
            misc_feature     27743..47303
                             /note="assembly_fragment"
            misc_feature     47404..73174
                             /note="assembly_fragment"
            misc_feature     73275..89069
                             /note="assembly_fragment"
            misc_feature     89170..122795
                             /note="assembly_fragment"
            misc_feature     122896..154841
                             /note="assembly_fragment"
                             clone_end:T7
                             vector_side:right"
  BASE COUNT   39087 a   38239 c   37968 g   38338 t   1209 others
  ORIGIN

  Query Match        4.5%;  Score 30;  DB 2;  Length 154841;
  Best Local Similarity  100.0%;  Pred. No. 0.00019;
  Matches  30;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

  Qy          229 CTGGCTCTCGCCCTCTCCCTCTCCCTCTCC 248
                  ||||||||||||||||||||||||||||||
  Db       124246 CTCGCTCTCGCCCTCTCCCTCTCCCTCTCC 124175

  RESULT 6
  AC004223/c
  LOCUS       AC004229           155736 bp    DNA    linear   HTG 13-JUN-2002
  DEFINITION  Homo sapiens chromosome 11 clone pUC57114, *** SEQUENCING IN
              PROGRESS ***, 29 unordered pieces.
  ACCESSION   AC004229
  VERSION     AC004229.2  GI:21406524
  KEYWORDS    HTG; HTGS_PHASE1.
  SOURCE      Homo sapiens (human)
    ORGANISM  Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE   1  (bases 1 to 155736)
    AUTHORS   Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
              Bumeiter,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
              Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
              Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
              Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
              Ward,T. and Wilson,R.
    TITLE     template
    JOURNAL   Unpublished
  REFERENCE   2  (bases 1 to 155736)
    AUTHORS   Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
              Bumeiter,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
              Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
              Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
```

TITLE   Schagemar,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
           Ward,T. and Wilson,R.
JCURNAL   Direct Submission
           Submitted (26-FEB-1998) Genome Science & Technology Center,
           University of Texas Southwester Medical Center, 5323 Harry Hines
           Blvd, Dallas, TX 75235-8591, USA
COMMENT   On Jun 13, 2002 this sequence version replaced gi:2911732.
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 29 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *   1   2239: contig of 2239 bp in length
           *   2240   2339: gap of 100 bp
           *   2340   4926: contig of 2587 bp in length
           *   4927   5026: gap of 100 bp
           *   5027   8003: contig of 2977 bp in length
           *   8004   8103: gap of 100 bp
           *   8104   10510: contig of 2407 bp in length
           *   10511   10610: gap of 100 bp
           *   10611   12841: contig of 2231 bp in length
           *   12842   12941: gap of 100 bp
           *   12942   16005: contig of 3064 bp in length
           *   16006   16105: gap of 100 bp
           *   16106   18268: contig of 2163 bp in length
           *   18269   18368: gap of 100 bp
           *   18369   21525: contig of 3157 bp in length
           *   21526   21625: gap of 100 bp
           *   21626   24492: contig of 2867 bp in length
           *   24493   24592: gap of 100 bp
           *   24593   27337: contig of 2745 bp in length
           *   27338   27437: gap of 100 bp
           *   27438   30341: contig of 2904 bp in length
           *   30342   30441: gap of 100 bp
           *   30442   34747: contig of 4306 bp in length
           *   34748   34847: gap of 100 bp
           *   34848   39221: contig of 4384 bp in length
           *   39222   39331: gap of 100 bp
           *   39332   43437: contig of 4106 bp in length
           *   43438   43537: gap of 100 bp
           *   43538   47029: contig of 3492 bp in length
           *   47030   47129: gap of 100 bp
           *   47130   53584: contig of 6455 bp in length
           *   53585   53685: gap of 100 bp
           *   53685   58032: contig of 4348 bp in length
           *   58033   58132: gap of 100 bp
           *   58133   63938: contig of 5806 bp in length
           *   63939   64038: gap of 100 bp
           *   64039   70229: contig of 6191 bp in length
           *   70230   70329: gap of 100 bp
           *   70330   76231: contig of 5902 bp in length
           *   76232   76331: gap of 100 bp
           *   76332   81290: contig of 4959 bp in length
           *   81291   81390: gap of 100 bp
           *   81391   88766: contig of 7376 bp in length
           *   88767   88866: gap of 100 bp
           *   88867   98893: contig of 10027 bp in length
           *   98894   98993: gap of 100 bp
           *   98994   106976: contig of 7993 bp in length
           *   106977   107076: gap of 100 bp
           *   107077   112608: contig of 5532 bp in length
           *   112609   112708: gap of 100 bp
           *   112709   121290: contig of 8582 bp in length
           *   121291   121390: gap of 100 bp
           *   121391   134770: contig of 13380 bp in length
           *   134771   134870: gap of 100 bp
           *   134871   147313: contig of 12443 bp in length
           *   147314   147413: gap of 100 bp
           *   147414   155736: contig of 8323 bp in length
FEATURES   Location/Qualifiers
           source   1..155736
                           /organism="Homo sapiens"
                           /mol_type="genomic DNA"
                           /db_xref="taxon:9606"
                           /chromosome="1"
                           /clone="pDJ571I4"
BASE COUNT   36015 a   38484 c   39960 g   38427 t   2870 others
ORIGIN

Query Match   4.5%;   Score 30;   DB 2;   Length 155736;
Best Local Similarity 100.0%;   Pred. No. 0.00019;
Matches   30; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Cy   219 CTCGCCTCGCCCTCTCCCTCTCCCTCTCC 248
              |||||||||||||||||||||||||||||
Db   93471 CTCGCCTCGCCCTCTCCCTCTCCCTCTCC 93442

RESULT 7
AX653891
LOCUS   AX653891   1092 bp   DNA   linear   PAT 22-MAR-2003
DEFINITION Sequence 3761 from Patent WO03000898.
ACCESSION   AX653891
VERSION   AX653891.1 GI:29156705
KEYWORDS
SOURCE   Oryza sativa
  ORGANISM Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   ;
  AUTHORS   Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
            Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
  TITLE   Plant genes involved in defense against pathogens
  JOURNAL   Patent: WO 03000898-A 3761 03-JAN-2003;
            Syngenta Participations AG (CH)
FEATURES   Location/Qualifiers
   source   1..1092
                   /organism="Oryza sativa"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:4530"
BASE COUNT   171 a   418 c   321 g   180 t   2 others
ORIGIN

Query Match   3.9%;   Score 26;   DB 6;   Length 1092;
Best Local Similarity 100.0%;   Pred. No. C.031;
Matches   26; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Cy   532 GCTCCTGCTCCTGCTCCTGCCCCCTC 557
              ||||||||||||||||||||||||||
Db   942 GCTCCTGCTCCTGCTCCTGCTCCCTC 967

RESULT 8
AL928648
LOCUS   AL928648   90637 bp   DNA   linear   ROD 19-MAR-2003
DEFINITION Mouse DNA sequence from clone RP23-379A6 on chromosome 4, complete
            sequence.
ACCESSION   AL928648
VERSION   AL928648.19 GI:29125387
KEYWORDS   HTG.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 90637)
  AUTHORS   Lovell,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Mar 19, 2003 this sequence version replaced gi:28810826.
            Sequence from the Mouse Genome Sequencing Consortium whole genome

shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-379A6 is
from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe.6.

FEATURES   Location/Qualifiers
     source   1..90837
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="4"
                     /clone="RP23-379A6"
                     /clone_lib="RPCI-23"

BASE COUNT   26550 a   19257 c   18758 g   26272 t
ORIGIN

Query Match   3.9%; Score 26; DB 10; Length 90837;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   223 CTCTCGCCCTCTCCCTCTCCCTCTCC 248
           ||||||||||||||||||||||||||
Db   7515 CTCTCGCCCTCTCCCTCTCCCTCTCC 7540

RESULT 9
AC133008
LOCUS   AC133008   120926 bp   DNA   linear   HTG 28-SEP-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 11 clone
            OSJNBb0026K20, *** SEQUENCING IN PROGRESS ***, 8 ordered pieces.
ACCESSION   AC133008
VERSION   AC133008.2 GI:23334734
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 120926)
AUTHORS   Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Bera,J., Jones,K.M.,
            Overton II,L., Tsitrin,T., Kim,M., Bera,J., Jin,S., Fadrosh,D.W.,
            Talion,L., Koo,H., Zismann,V., Hsiao,J., Blunt,S., Vanaken,S.,
            Riedmuller,S.B., Utterback,T., Feldblyum,T., Yang,Q., Haas,B.,
            Suh,B., Peterson,J., Quackenbush,J., White,O., Salzberg,S. and
            Fraser,C
TITLE   Oryza sativa ssp. japonica cv. Nipponbare OSJNBb0026K20 BAC genomic

JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 120926)
AUTHORS   Buell,R.
TITLE   Direct Submission
JOURNAL   Submitted (05-SEP-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   3 (bases 1 to 120926)
AUTHORS   Buell,R.
TITLE   Direct Submission
JOURNAL   Submitted (28-SEP-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT   On Sep 27, 2002 this sequence version replaced gi:22726029.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *   1: contig of 342 bp in length
            * 3442: gap of unknown length
            * 3543: contig of 2318 bp in length
            * 5860: gap of unknown length
            * 5861: contig of unknown length
            * 5960: gap of unknown length
            * 18145: contig of 12185 bp in length
            * 18146: gap of unknown length
            * 18245: contig of 54108 bp in length
            * 72353: gap of unknown length
            * 72354: contig of unknown length
            * 72453: gap of unknown length
            * 72454: contig of 1609 bp in length
            * 88544: gap of unknown length
            * 88545: contig of unknown length
            * 88644: gap of unknown length
            * 88645: contig of 20926 bp in length
            *109570: gap of unknown length
            *109571: contig of 2100 bp in length
            *109671: gap of unknown length
            *111770: contig of unknown length
            *111771: gap of unknown length
            *111870: contig of unknown length
            *111871: contig of 9056 bp in length
            *120926:

FEATURES   Location/Qualifiers
     source   1..120926
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="11"
                     /clone="OSJNBb0026K20"
                     /note="japonica cultivar-group"

BASE COUNT   34060 a   26488 c   25743 g   33915 t   720 others
ORIGIN

Query Match   3.9%; Score 26; DB 2; Length 120926;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   532 GCTCCTGCTCCTCCTGCTCCTCCCTC 557
           ||||||||||||||||||||||||||
Db 78382 GCTCCTGCTCCTCCTGCTCCTCCCTC 78407

RESULT 10
AC023549
LOCUS   AC023549   155185 bp   DNA   linear   HTG 10-SEP-2000
DEFINITION Homo sapiens clone RP11-60C17, WORKING DRAFT SEQUENCE, 17 unordered
            pieces.
ACCESSION   AC023549
VERSION   AC023549.3 GI:10047834
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 155185)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE   Homo sapiens chromosome, clone RP11-60C17

```
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 155185)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Browe,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Gaiagar,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., Mckernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Miranda,J., Mlenga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oivar,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,A., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Triglio,J., Vassiliev,H., Viei,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 10, 2000 this sequence version replaced gi:7331469.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            --------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            --------------- Project Information
            Center project name: L6394
            Center clone name: 60_C17
            --------------- Summary Statistics
            Sequencing vector: M13; X77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 144294 bases at least Q40
            Consensus quality: 148497 bases at least Q30
            Consensus quality: 150798 bases at least Q20
            Insert size: 148000; agarose-fp
            Insert size: 153585; sum-of-contigs
            Quality coverage: 5.0 in Q20 bases; agarose-fp
            Quality coverage: 4.8 in Q20 bases; sum-of-contigs
            --------------------------------------
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 17 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *      1  1127: contig of 1127 bp in length
          *   1128  1227: gap of 100 bp
          *   1228  2232: contig of 1005 bp in length
          *   2233  2332: gap of 100 bp
          *   2333  3444: contig of 1112 bp in length
          *   3445  3544: gap of 100 bp
          *   3545  5242: contig of 1698 bp in length
          *   5243  5342: gap of 100 bp
          *   5343  8059: contig of 2717 bp in length
          *   8060  8159: gap of 100 bp
          *   8160 11253: contig of 3094 bp in length
          *  11254 11353: gap of 100 bp
          *  11354 14824: contig of 3471 bp in length
          *  14825 14924: gap of 100 bp
          *  14925 49854: contig of 34930 bp in length
```

```
          *  49855  49954: gap of 100 bp
          *  49955  53870: contig of 3916 bp in length
          *  53871  53970: gap of 100 bp
          *  53971  59531: contig of 5561 bp in length
          *  59532  59631: gap of 100 bp
          *  59632  65752: contig of 6121 bp in length
          *  65753  65852: gap of 100 bp
          *  65853  71898: contig of 6046 bp in length
          *  71899  71998: gap of 100 bp
          *  71999  79996: contig of 7998 bp in length
          *  79997  80096: gap of 100 bp
          *  80097  97717: contig of 17621 bp in length
          *  97718  97817: gap of 100 bp
          *  97818 122075: contig of 24258 bp in length
          * 122076 122175: gap of 100 bp
          * 122176 152445: contig of 30270 bp in length
          * 152446 152545: gap of 100 bp
          * 152546 155185: contig of 2640 bp in length

FEATURES             Location/Qualifiers
     source          1..155185
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RP11-60C17"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1127
                     /note="assembly_fragment"
     misc_feature    1228..2232
                     /note="assembly_fragment"
     misc_feature    2333..3444
                     /note="assembly_fragment"
     misc_feature    3545..5242
                     /note="assembly_fragment"
     misc_feature    5343..8059
                     /note="assembly_fragment"
     misc_feature    8160..11253
                     /note="assembly_fragment"
     misc_feature    11354..14824
                     /note="assembly_fragment"
     misc_feature    14925..49854
                     /note="assembly_fragment"
     misc_feature    49955..53870
                     /note="assembly_fragment"
     misc_feature    59632..65752
                     /note="assembly_fragment"
     misc_feature    65853..71898
                     /note="assembly_fragment"
     misc_feature    71999..79996
                     /note="assembly_fragment"
     misc_feature    80097..97717
                     /note="assembly_fragment"
     misc_feature    97818..122075
                     /note="assembly_fragment"
     misc_feature    122176..152445
                     /note="assembly_fragment"
     misc_feature    152546..155185
                     /note="assembly_fragment"
                     clone_end:T7
                     vector_side:right"

BASE COUNT     47375 a    29006 c    28938 g    48262 t     1604 others
ORIGIN
Query Match         3.9%;  Score 26;  DB 2;  Length 155185;
Best Local Similarity  100.0%;  Pred. No. 0.026;
Matches 26;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    223  CTCTCGCCCTCTCCCCTCGCTCTCTCC 248
               |||||||||||||||||||||||||||
Db  54935  CTCTCGCCCTCTCCCCTCGCTCTCTCC 54960
```

```
RESULT 11
ACCESSION   AC022931
LOCUS       AC022931            157834 bp    DNA     linear   HTG 16-MAR-2000
DEFINITION  Homo sapiens chromosome 5 clone RP11-356D23 map 5, WORKING DRAFT
            SEQUENCE, 17 unordered pieces.
ACCESSION   AC022931
VERSION     AC022931.3  GI:7249435
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157834)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 5, clone RP11-356D23
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 157834)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgaiter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Co-zymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Leboczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meidrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Normar,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.G.,
            Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 16, 2000 this sequence version replaced gi:7008859.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
            Center project name: L6343
            Center clone name: 356_D_23
            -------------- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 148896 bases at least Q40
            Consensus quality: 153119 bases at least Q30
            Consensus quality: 154676 bases at least Q20
            Insert size: 160000; agarose-fp
            Insert size: 156234; sum-of-contigs
            Quality coverage: 4.8 in Q20 bases; agarose-fp
            Quality coverage: 4.9 in Q20 bases; sum-of-contigs
            --------------
            NOTE: This is a 'working draft' sequence. It currently
            consists of 17 contigs. The true order of the pieces
            is not known and their order in this sequence record is
            arbitrary. Gaps between the contigs are represented as
            runs of N, but the exact sizes of the gaps are unknown.
            This record will be updated with the finished sequence
            as soon as it is available and the accession number will
            be preserved.
                1    1404: contig of 1404 bp in length
             1405    1504: gap of 100 bp
             1505    4290: contig of 2786 bp in length
             4291    4390: gap of 100 bp
```

```
            4391    7101: contig of 2711 bp in length
            7102    7201: gap of 100 bp
            7202   10425: contig of 3224 bp in length
           10426   10525: gap of 100 bp
           10526   14004: contig of 3479 bp in length
           14005   14104: gap of 100 bp
           14105   18152: contig of 4048 bp in length
           18153   18252: gap of 100 bp
           18253   22230: contig of 3978 bp in length
           22231   22330: gap of 100 bp
           22331   28231: contig of 5901 bp in length
           28232   28331: gap of 100 bp
           28332   34381: contig of 6050 bp in length
           34382   34481: gap of 100 bp
           34482   40840: contig of 6359 bp in length
           40841   40940: gap of 100 bp
           40941   48630: contig of 7690 bp in length
           48631   48730: gap of 100 bp
           48731   60384: contig of 11654 bp in length
           60385   60484: gap of 100 bp
           60485   72266: contig of 11782 bp in length
           72267   72366: gap of 100 bp
           72367   85070: contig of 12704 bp in length
           85071   85170: gap of 100 bp
           85171  100748: contig of 15578 bp in length
          100749  100848: gap of 100 bp
          100849  121955: contig of 21107 bp in length
          121956  122055: gap of 100 bp
          122056  157834: contig of 35779 bp in length

FEATURES             Location/Qualifiers
     source          1..157834
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5"
                     /clone="RP11-356D23"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..464
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    1505..4290
     misc_feature    4391..7101
                     /note="assembly_fragment"
     misc_feature    7202..10425
                     /note="assembly_fragment"
     misc_feature    10526..14004
                     /note="assembly_fragment"
     misc_feature    14105..18152
                     /note="assembly_fragment"
     misc_feature    18253..22230
                     /note="assembly_fragment"
     misc_feature    22331..28231
                     /note="assembly_fragment"
     misc_feature    28332..34381
                     /note="assembly_fragment"
     misc_feature    34482..40840
                     /note="assembly_fragment"
     misc_feature    40941..48630
                     /note="assembly_fragment"
     misc_feature    48731..60384
                     /note="assembly_fragment"
     misc_feature    60485..72266
                     /note="assembly_fragment"
     misc_feature    72367..85070
                     /note="assembly_fragment"
     misc_feature    85171..100748
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    100849..121955
```

```
misc_feature            /note="assembly_fragment"
                        122056..157834
                        /note="assembly_fragment"
BASE COUNT    46053 a  31958 c  29611 g  48607 t   1605 others
ORIGIN

Query Match
Best Local Similarity 100.0%;  Pred.No. 0.026;  Length 157834;
Matches 26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     223 CTCTCGCCCTCTCCCTCTCCCTCTCC 248
           ||||||||||||||||||||||||||
Db    6178 CTCTCGCCCTCTCCCTCTCCCTCTCC 6203

RESULT 12
AC114969
LOCUS       AC114969           164526 bp    DNA     linear   PRI 01-JUN-2002
DEFINITION  Homo sapiens chromosome 5 clone RP11-356D23, complete sequence.
ACCESSION   AC114969
VERSION     AC114969.2  GI:21306652
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 164526)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE       Direct Submission
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 164526)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 164526)
AUTHORS     DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE       Direct Submission
JOURNAL     Submitted (01-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jun 1, 2002 this sequence version replaced gi:19424430.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.8% of sequence;
            Estimated Total Number of Errors is 0.2.
FEATURES             Location/Qualifiers
     source          1..164526
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-356D23"
BASE COUNT    52577 a  32245 c  32724 g  46980 t
ORIGIN

Query Match
Best Local Similarity 100.0%;  Pred.No. 0.026;  Length 164526;
Matches 26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     223 CTCTCGCCCTCTCCCTCTCCCTCTCC 248
           ||||||||||||||||||||||||||
Db   94483 CTCTCGCCCTCTCCCTCTCCCTCTCC 94508

RESULT 13
BX293990/c
LOCUS       BX293990           183011 bp    DNA     linear   HTG 02-MAY-2003
DEFINITION  Mus musculus chromosome 2 clone RP24-222N16, *** SEQUENCING IN
            PROGRESS ***.
ACCESSION   BX293990
VERSION     BX293990.5  GI:30348793
```

```
KEYWORDS    HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 183011)
AUTHORS     Johnson.C.
TITLE       Direct Submission
JOURNAL     Submitted (01-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
COMMENT     On May 2, 2003 this sequence version replaced gi:30230876.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence. Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bN222N16
            -------------- Summary Statistics
            Chemistry: Dye-terminator; 100% of reads
            Assembly program: XGAP4; version 4.5
            Consensus quality: 182293 bases at least Q40
            Consensus quality: 182531 bases at least Q30
            Consensus quality: 182695 bases at least Q20
            Insert size: 183011; sum-of-contigs
            Insert size: 169280; 7.4% error; agarose-fp
            Quality coverage: 8.90x in Q20 bases; sum-of-contigs Quality
            coverage: 9.70x in Q20 bases; agarose-fp
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 1 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submitter.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *    1..183011: contig of 183011 bp in length.
FEATURES             Location/Qualifiers
     source          1..183011
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP24-222N16"
                     /clone_lib="RPCI-24"
     misc_feature    1..183011
                     /note="assembly_fragment:02752"
                     clone_end:T7
                     vector_side:left"
BASE COUNT    50069 a  41722 c  40049 g  51571 t
ORIGIN

Query Match
Best Local Similarity 100.0%;  Pred. No. 0.025;  Length 183011;
Matches 26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     223 CTCTCGCCCTCTCCCTCTCCCTCTCC 248
           ||||||||||||||||||||||||||
Db  162078 CTCTCGCCCTCTCCCTCTCCCTCTCC 162053

RESULT 14
AC129216/c
LOCUS       AC129216           190338 bp    DNA     linear   ROD 23-APR-2003
DEFINITION  Mus musculus chromosome 14 clone RP24-446G11, complete sequence.
ACCESSION   AC129216
```

VERSION   AC129216.4   GI:30061471
KEYWORDS   HTG.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 190338)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE   The sequence of Mus musculus clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 190338)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (27-JUL-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3 (bases 1 to 190338)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (20-MAR-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4 (bases 1 to 190338)
AUTHORS   McPherson,J.D. and Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (23-APR-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT   On Apr 23, 2003 this sequence version replaced gi:29126529.

            ----------------------- Genome Center -----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ------------------ Project Information ------------------
            Center project name: M_BB0446C11

FEATURES   Location/Qualifiers
     source   1..190338
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /clone="RP24-446C11"

BASE COUNT   52061 a 38827 c 41432 g 58018 t
ORIGIN

Query Match   3.9%; Score 26; DB 10; Length 190338;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches   26; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   223 CTCTCGCCCTCTCCCTCCCTCTCC 248
          ||||||||||||||||||||||||
Db 141919 CTCTCGCCCTCTCCCTCCCTCTCC 141894

RESULT 15
AL928914/c
LOCUS   AL928914   199873 bp   DNA   linear   ROD 01-JUN-2003
DEFINITION Mouse DNA sequence from clone RP23-399J8 on chromosome 2, complete
            sequence.
ACCESSION   AL928914
VERSION   AL928914.15 GI:31135573
KEYWORDS   HTG.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 199873)
AUTHORS   Humphries,M.
TITLE   Direct Submission
JOURNAL   Submitted (30-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT   On Jun 2, 2003 this sequence version replaced gi:29888620.

Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
------------ Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP. Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-399J8 is
from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES   Location/Qualifiers
     source   1..199873
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="2"
                     /clone="RP23-399J8"
                     /clone_lib="RPCI-23"

BASE COUNT   50089 a 47425 c 46975 g 55384 t
ORIGIN

Query Match   3.9%; Score 26; DB 10; Length 199873;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches   26; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   223 CTCTCGCCCTCTCCCTCCCTCTCC 248
          ||||||||||||||||||||||||
Db 89873 CTCTCGCCCTCTCCCTCCCTCTCC 89848

Search completed: November 17, 2003, 21:16:26
Job time : 2571 secs